# A Review on Suspicious Activity Prediction Based on Deep Learning Model

Bhondave Ajitkumar Balkrishna[1], Prof. Rahane Divya[2], Kakade Nilesh Samadhan[3]
[1,3]*Student, Department of Artificial Intelligence &Data Science Engineering Sahyadri Valley College of Engineering & Technology Rajuri, India. Savitribai Phule Pune University*
[2]*HOD, Department of Artificial Intelligence &Data Science Engineering Sahyadri Valley College of Engineering &Technology Rajuri, India. Savitribai Phule Pune University*

***Abstract*—Suspicious activity prediction is vital for enhancing security and surveillance systems by identifying potential threats in real time. This study implements a deep learning-based approach using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to analyze video streams and predict abnormal human behavior. OpenCV is utilized for video processing, while TensorFlow handles model training and inference. The system effectively learns spatial and temporal features to distinguish between normal and suspicious activities. Performance is evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results demonstrate that the proposed framework accurately predicts suspicious activities in real time, offering an efficient and reliable solution for automated surveillance applications.**

***Index Terms*—Deep learning, computer vision, suspicious activity detection, real-time surveillance, human behavior analysis, convolutional neural networks (CNN), activity recognition, anomaly detection, OpenCV, artificial intelligence in security.**

## I. INTRODUCTION

In the field of public safety and surveillance, detecting suspicious activities plays a crucial role in preventing potential crimes and ensuring the security of individuals and property. Traditional security monitoring methods rely heavily on human supervision, which can be labour-intensive, inconsistent, and prone to error due to fatigue or inattention. Recent advancements in artificial intelligence, particularly in deep learning and computer vision, have enabled the development of intelligent systems capable of automatically identifying unusual or suspicious behavior from video feeds.

Convolutional Neural Networks (CNNs) and other deep learning models have demonstrated exceptional performance in visual recognition, motion analysis, and activity classification tasks. By leveraging these models, it becomes possible to analyze video frames in real time, detect anomalies in human movements, and flag potential threats or suspicious actions such as loitering, fighting, or unauthorized access.

This project proposes an automated suspicious activity prediction system that processes live or recorded video feeds using deep learning algorithms. The input video frames undergo several preprocessing stages such as resizing, normalization, and feature extraction to enhance detection accuracy.

The trained model then analyzes these frames to identify and classify activities based on behavioral patterns.

By integrating this intelligent detection framework with modern surveillance infrastructure, security operations can achieve higher accuracy, reduced response time, and lower dependence on human monitoring. Such systems can be applied in various domains including public safety, smart cities, transportation hubs, and private security, thereby contributing to a safer and more responsive environment. Implementing.

## II. DISCUSSION

In recent years, numerous intelligent surveillance and activity recognition systems have been developed to enhance public safety and automate threat detection. However, many existing applications still face challenges such as low detection accuracy in complex environments, high computational costs, and limited

adaptability to different activity types or lighting conditions. These limitations often reduce their reliability in real-world scenarios like public areas, transportation hubs, and private security systems.

The proposed Suspicious Activity Prediction System overcomes these challenges by leveraging Deep Learning architectures, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), for efficient spatiotemporal analysis of human motion and behavior. The system detects and classifies unusual activities—such as fighting, loitering, or trespassing—directly from video frames in real time. By applying transfer learning on pre-trained models such as YOLO or LSTM-based action classifiers, the system achieves high performance even with limited domain-specific datasets, ensuring better generalization across diverse environments.

The development follows an Agile methodology, allowing iterative testing and continuous integration of user feedback from surveillance operators and stakeholders. This ensures the software remains flexible, scalable, and aligned with practical needs. The implementation utilizes OpenCV for video stream handling and TensorFlow or PyTorch for deep learning inference, supporting both cloud-based and edge-device deployment.

## III. LITERATURE REVIEW

Suspicious activity prediction, an important aspect of intelligent surveillance, has gained considerable attention due to its wide range of applications in security, public safety, and smart monitoring systems. Early research in this field primarily relied on traditional machine learning techniques and handcrafted features, such as motion trajectories, optical flow, and background subtraction, to detect abnormal behavior. While these approaches provided useful insights, they often struggled to handle complex environments, varying lighting conditions, and diverse human actions.

The emergence of deep learning has brought a major transformation in the field of activity recognition and prediction. Convolutional Neural Networks (CNNs) have proven highly effective in extracting spatial features from video frames, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks capture temporal dependencies across sequences. The combination of CNN and LSTM architectures has demonstrated superior performance in modeling human motion and identifying suspicious or abnormal behavior in dynamic scenes. Recent advancements also include the use of 3D CNNs, Autoencoders, and Transformer-based architectures for more accurate spatio-temporal feature extraction. These deep learning models enable real-time analysis and prediction, making them suitable for modern surveillance applications where early detection of suspicious activity is critical. Consequently, deep learning-based methods have become the foundation for intelligent security systems, outperforming traditional techniques in both accuracy and adaptability

## IV. SYSTEM IMPLEMENTATION

Implementing a system for *Suspicious Activity Prediction based on Deep Learning* involves several systematic stages. Below is a high-level overview of the key processes involved:

1. Dataset Compilation:
Gather a comprehensive dataset containing video clips or image sequences that represent both normal and suspicious human activities. Publicly available datasets such as *UCF Crime* or *UCSD Pedestrian Anomaly Dataset* can be used. Each video segment should be labeled according to the type of activity it depicts.

2. Data Preprocessing:
Extract frames from videos and resize them to a uniform resolution compatible with the deep learning model. Normalize pixel values, remove noise, and apply data augmentation techniques such as rotation, scaling, and brightness adjustment to enhance the model's generalization capability.

3. Feature Extraction:
Utilize Convolutional Neural Networks (CNNs) to extract spatial features from each video frame. These features represent visual information such as human posture, motion, and background context, which are essential for distinguishing suspicious activities.

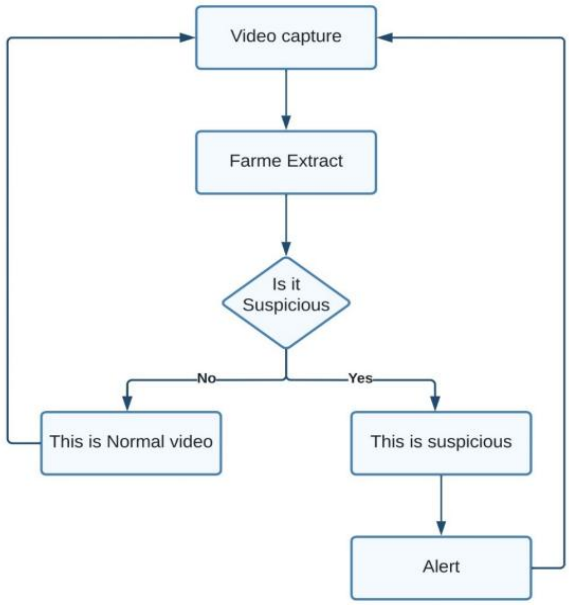
4. Temporal Modeling:
Integrate Long Short-Term Memory (LSTM) networks to capture temporal dependencies between

consecutive frames. The LSTM analyzes the sequence of actions over time, enabling the system to detect unusual motion patterns that may indicate suspicious behavior.

5. Model Training:
Combine the CNN and LSTM architectures into a unified deep learning model. Train the model using the labeled dataset, optimizing parameters through backpropagation. Use validation data to fine-tune hyperparameters and prevent overfitting.

6. Performance Evaluation:
Evaluate the trained model using a separate testing dataset. Key performance metrics include accuracy, precision, recall, and F1-score, which measure the model's reliability in identifying suspicious activities.

7. System Integration and Deployment:
Develop a real-time surveillance application that integrates the trained model. The system should process live video streams, predict suspicious activities, and trigger alerts when abnormal behavior is detected. Continuous monitoring and periodic retraining ensure consistent system accuracy and efficiency.




## V. SYSTEM ARCHITECTURE

The system is divided into 2 phases: training and deployment. During the training phase, the ResNet model is trained with our custom dataset. We divide the dataset into training and validation sets. Validation set contained 20% of the images which were randomly chosen from the dataset.
After the training phase, the model is deployed on computer systems used by the security teams in public places. Our system is a desktop application which can take as input live feed from a camera or an already stored video from the computer. This video is then preprocessed (which involves breaking the video into frames) and then fed into the ResNet-50 model. The model outputs if the video contains any suspicious activity or not. If a suspicious activity is detected in the video, the model immediately generates an alert on the system and also sends an email alert on the registered email address along with pictures of the video where suspicious activity is ongoing.

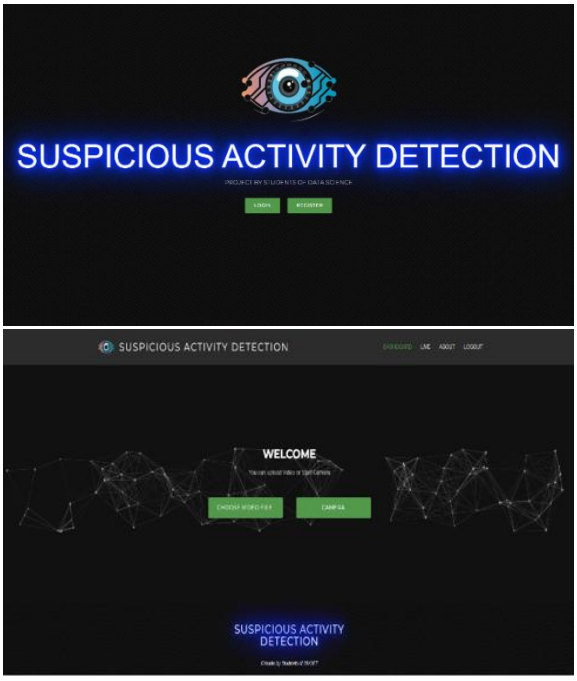



Fig 1: project screenshot

## VI. SYSTEM REQUIREMENTS

*Hardware Interfaces:*
RAM: 8 GB As we are using Machine Learning Algorithm and Various High Level Libraries Laptop RAM minimum required is 8 GB.
Hard Disk: 40 GB Data Set of CT Scan images is to be used hence minimum 40 GB Hard Disk memory is required.
Processor: Intel i5 Processor
IDE: pycharm

Coding Language: python
Operating System: Windows 10 XX 4.1.3

*Software Interfaces:*
Operating System: Windows 10
IDE: pyCharm
Programming Language: python

## VII. SCOPE

The future of suspicious activity prediction based on deep learning holds immense potential, driven by rapid advancements in artificial intelligence and computer vision technologies. Future research is expected to focus on improving model accuracy, real-time performance, and contextual understanding of human behavior in complex environments.

With the integration of advanced architectures such as 3D Convolutional Neural Networks (3D-CNNs), Transformers, and Graph Neural Networks (GNNs), models will be able to capture more detailed spatio-temporal relationships, leading to more precise activity recognition. The incorporation of multimodal data—including audio, motion sensors, and thermal imagery—will further enhance system reliability in diverse surveillance conditions.

Additionally, the implementation of edge computing and cloud-based solutions will enable faster and more efficient processing of live video streams, allowing real-time suspicious activity detection in large-scale surveillance systems. Transfer learning and continual learning techniques will also play a key role in adapting models to new environments without retraining from scratch.

In the long term, such intelligent systems can be integrated into smart cities, public transportation, and critical infrastructure, significantly improving public safety and proactive threat prevention.

## VIII. CONCLUSION

suspicious activity prediction based on deep learning represents a modern and effective approach to enhancing surveillance and public safety systems. By leveraging the combined strengths of Convolutional Neural Networks (CNNs) for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal analysis, the system can accurately identify and predict abnormal human behaviors in real time. This integration enables automated monitoring with improved precision, reducing the dependency on manual observation.

Although the approach demonstrates strong performance, challenges such as limited datasets, computational demands, and interpretability issues still persist. Continuous research in model optimization, real-time processing, and adaptive learning is expected to address these limitations.

Overall, deep learning-based suspicious activity prediction systems hold significant potential for applications in public security, smart cities, and intelligent surveillance, paving the way for safer and more proactive environments through intelligent video analytics.

## REFERENCES

[1] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, “OpenPose: Realtime multi-person 2D pose estimation using part affinity fields,” *arXiv preprint arXiv:1812.08008*, 2018.

[2] A. Kendall, M. Grimes, and R. Cipolla, “PoseNet: A convolutional network for real-time 6-DOF camera relocalization,” arXiv preprint arXiv:1505.07427, 2015.

[3] J. Howard and S. Gugger, Deep Learning for Coders with fastai and PyTorch: AI Applications Without a PhD. Sebastopol, CA, USA: O’Reilly Media, 2020.

[4] J. Howard and S. Gugger, “fastai: A layered API for deep learning,” Information, vol. 11, no. 2, p. 108, 2020.

[5] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., “PyTorch: An imperative style, high-performance deep learning library,” in Advances in Neural Information Processing Systems (NeurIPS 2019), vol. 32, Vancouver, BC, Canada, 2019, pp. 8024–8035.